

D. J.  
#5 2-14-02104  
Priority Papers

日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

J1033 U.S. PTO  
09/891717  
06/26/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2000年 6月26日

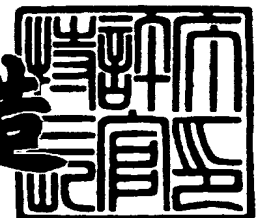
出 願 番 号  
Application Number: 特願2000-191573

出 願 人  
Applicant (s): インターナショナル・ビジネス・マシーンズ・コーポレーション

2000年 8月18日

特許庁長官  
Commissioner,  
Patent Office

及 川 耕 造



出証番号 出証特2000-3065322

【書類名】 特許願

【整理番号】 JP9000104

【提出日】 平成12年 6月26日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 17/20

【発明者】

    【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4    日本アイ・ビー・エム株式会社 大和事業所内

    【氏名】 酒井 秀雄

【特許出願人】

    【識別番号】 390009531

    【氏名又は名称】 インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

    【識別番号】 100086243

    【弁理士】

    【氏名又は名称】 坂口 博

【復代理人】

    【識別番号】 100104880

    【弁理士】

    【氏名又は名称】 古部 次郎

【選任した代理人】

    【識別番号】 100091568

    【弁理士】

    【氏名又は名称】 市位 嘉宏

【選任した復代理人】

    【識別番号】 100100077

    【弁理士】

    【氏名又は名称】 大場 充

【手数料の表示】

【予納台帳番号】 081504

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9706050

【包括委任状番号】 9704733

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成システム、音声合成方法、サーバ、記憶媒体、プログラム伝送装置、音声合成データ記憶媒体、音声出力機器

【特許請求の範囲】

【請求項 1】 顧客とサービス提供者との間でネットワークを介してなされる音声合成システムであって、

前記顧客に対し選択可能に提示された話者の中から特定の話者を選択するとともに、音声合成すべき文章を指定する前記顧客の端末と、

前記顧客が前記端末で指定した前記音声合成すべき文章を、前記特定の話者の音声特徴データを用いて音声合成する前記サービス提供者のサーバと、を含むことを特徴とする音声合成システム。

【請求項 2】 前記サービス提供者の前記サーバは、音声合成することにより生成した音声合成データを、前記ネットワークを介して前記顧客の前記端末に受け渡すことを特徴とする請求項 1 記載の音声合成システム。

【請求項 3】 前記サービス提供者の前記サーバは、前記顧客に対し取引番号を付与し、

前記顧客の前記端末から前記取引番号が提示されたときに、生成した前記音声合成データを前記顧客の前記端末に受け渡すことを特徴とする請求項 2 記載の音声合成システム。

【請求項 4】 複数の話者の音声特徴データを保有したサービス提供者と顧客との間におけるネットワークを介しての音声合成方法であって、

前記サービス提供者は、複数の前記話者のリストを、前記ネットワークを介して外部に提示し、

前記顧客は、提示された前記リストから選択した話者と、音声合成すべき文章とを、前記ネットワークを介して前記サービス提供者に伝え、

前記サービス提供者は、前記顧客の選択した前記話者の音声特徴データに基づき、前記音声合成すべき文章を音声合成することを特徴とする音声合成方法。

【請求項 5】 前記サービス提供者は、前記顧客に対して、音声合成して生成する音声合成データについて課金し、

前記サービス提供者は、前記顧客側から前記課金した代金を受け取った後に前記音声合成データを前記顧客に受け渡すことを特徴とする請求項 4 記載の音声合成方法。

【請求項 6】 前記サービス提供者は、当該サービス提供者が保有する音声特徴データの権利を有する者に対し、音声合成データの生成に応じた代金を支払うことを特徴とする請求項 4 記載の音声合成方法。

【請求項 7】 前記サービス提供者は、音声合成して生成した音声合成データを前記顧客に受け渡し、

当該顧客は、受け渡された前記音声合成データを、当該音声合成データに基づいた音声再生可能な機器に入力することを特徴とする請求項 4 記載の音声合成方法。

【請求項 8】 前記サービス提供者は、前記顧客に対し、音声合成データを入力すべき機器のリストを、前記話者のリストとともに提示し、

前記顧客は、前記機器のリストから選択した機器を前記ネットワークを介して前記サービス提供者に伝え、

前記サービス提供者は、前記顧客の選択した前記話者の音声特徴データに基づく音声合成データを生成するとともに、生成した当該音声合成データを前記顧客が選択した前記機器に入力することを特徴とする請求項 4 記載の音声合成方法。

【請求項 9】 ネットワーク上に接続された顧客の要求に応じて音声合成するサーバであって、

話者の音声を分析して得られた音声特徴データを格納する音声特徴データ格納部と、

前記ネットワークを介して、音声合成すべき文章と話者の指定を含む顧客からの要求を受け付ける要求受付部と、

前記要求受付部を介して入力された顧客からの要求に応じ、前記音声特徴データ格納部に格納されかつ前記話者に対応した音声特徴データに基づいて、前記文章を音声合成する音声合成データ生成部と、  
を備えることを特徴とするサーバ。

【請求項 10】 前記音声特徴データ格納部は、音声特徴データとして、話

者毎の声質データと韻律データとを格納することを特徴とする請求項 9 記載のサーバ。

【請求項 1 1】 前記顧客によってなされた要求に基づき、生成する前記音声合成データの価格を設定する価格設定部とをさらに備えることを特徴とする請求項 9 記載のサーバ。

【請求項 1 2】 コンピュータに実行させるプログラムを当該コンピュータの入力手段が読み取り可能に記憶した記憶媒体において、

前記プログラムは、外部から音声合成データを生成するための要求を受け付ける処理と、

受け付けた前記要求に対し、取引番号を生成して出力する処理と、

前記取引番号が入力されたときに、受け付けた前記要求に応じた音声合成データを出力する処理と、を前記コンピュータに実行させることを特徴とする記憶媒体。

【請求項 1 3】 前記プログラムは、生成した前記音声合成データに、当該音声合成データの内容を証明する認証データを添付する処理と、をさらに前記コンピュータに実行させることを特徴とする請求項 1 2 記載の記憶媒体。

【請求項 1 4】 コンピュータに実行させるプログラムを、当該コンピュータの入力手段が読み取り可能に記憶した記憶媒体において、

前記プログラムは、音声合成すべき文章の指定と話者の指定を含む外部からの要求を受け付ける処理と、

受け付けた要求に応じ、前記話者に対応した音声特徴データに基づいて前記文章を音声合成する処理と、を前記コンピュータに実行させることを特徴とする記憶媒体。

【請求項 1 5】 コンピュータに、当該コンピュータに格納している複数の音声特徴データのリストを顧客に対して出力する処理と、前記顧客が入力した文章を当該顧客が前記リストから選択した音声特徴データに基づいて音声合成した音声合成データを、前記顧客に出力する処理をコンピュータに実行させるプログラムが記憶された記憶手段と、

前記記憶手段から前記プログラムを読み出して当該プログラムを送信する送信

手段とを備えたことを特徴とするプログラム伝送装置。

【請求項 1 6】 ネットワークを介してサービス提供者に接続された顧客が、話者と文章とを前記サービス提供者に指定し、当該サービス提供者が前記顧客の指定に応じて生成した音声合成データを格納したことを特徴とする音声合成データ記憶媒体。

【請求項 1 7】 複数の話者の音声データを保有するサービス提供者に対しネットワークを介して指定された話者および文章に基づき、前記サービス提供者が生成した音声合成データを記憶する記憶部と、

当該記憶部で記憶した前記音声合成データに基づく音声を出力する音声出力部と、を備えることを特徴とする音声出力機器。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、特定キャラクタの音声を合成して生成する音声合成データを、ネットワークを介して取引することを可能とする音声合成システム、音声合成方法、サーバ、記憶媒体、プログラム伝送装置、音声合成データ記憶媒体、音声出力機器に関する。

【0 0 0 2】

【従来の技術】

従来より、歌手、政治家等の有名人、テレビ番組や映画の登場人物等、特定キャラクタの音声を、玩具、目覚し時計、携帯型電話端末等、種々の機器に組み込んだ製品が提供されている。このような製品では、所定の操作を行ったときに、特定キャラクタの音声によるメッセージが出力され、これにより製品の付加価値が高められている。

【0 0 0 3】

【発明が解決しようとする課題】

しかしながら、従来は、機器メーカー側で、特定キャラクタの音声による、予め決められた文章の音声データを機器に収めているに過ぎず、購入者(顧客)の好みでメッセージの文章を設定することはできなかった。

【 0 0 0 4 】

また、近年の音声合成技術の進歩により、特定キャラクタの音声の声質・韻律等の特徴をデータ化し、そのデータから、入力した文章に基づいて、特定キャラクタの音声に酷似した音声合成メッセージを生成することが可能となっている。

このような技術を機器メーカー側で用いる場合には、特定キャラクタに対して著作権料の支払を契約によって明確にすることができるので特に問題はない。しかし、上記したような技術を、例えばソフトウェアの形態でユーザ(購入者)に提供(販売)すると、ユーザ側において音声合成メッセージを自由に生成することが可能となり、このような場合には、特定キャラクタに対する著作権料の支払が不明確になるという側面もある。

【 0 0 0 5 】

本発明は、このような技術的課題に基づいて、顧客の好みに応じた音声合成メッセージを提供する音声合成システム、音声合成方法、サーバ、記憶媒体、プログラム伝送装置、音声合成データ記憶媒体、音声出力機器を提供することを主たる目的とする。

また、他の目的は、特定キャラクタの音声を用いる場合、著作権料を確実に支払い、特定キャラクタの権利を保護することにある。

【 0 0 0 6 】

【課題を解決するための手段】

かかる目的のもと、本発明の音声合成システムは、顧客に対し選択可能に提示された話者の中から特定の話者を選択するとともに、音声合成すべき文章を指定する顧客の端末と、顧客が指定した音声合成すべき文章を、特定の話者の音声特徴データを用いて音声合成するサービス提供者のサーバと、を含むことを特徴としている。このような構成により、顧客は、歌手、政治家等の有名人、テレビ番組や映画の登場人物等、自分の希望する話者の音声による自由なメッセージや歌等の音声合成データをオーダして入手することができる。そして入手した音声合成データにより、例えば目覚し時計の目覚しアラーム、携帯型電話端末の呼び出し音(メッセージ)や留守応答メッセージ、カーナビゲーション装置におけるガイドンスメッセージ等を自分の好みに応じて設定することが可能となるのである。



## 【 0 0 0 7 】

更に、サービス提供者のサーバは、顧客に対し取引番号を付与し、顧客の端末から取引番号が提示されたときに、生成した音声合成データを受け渡すことを特徴とすれば、オーダした顧客のみに音声合成データを受け渡すことができる。つまり、生成した音声合成データは、顧客以外の手に渡ることのない唯一無二のものとなるのである。

## 【 0 0 0 8 】

本発明の音声合成方法は、サービス提供者が、複数の話者のリストをネットワークを介して外部に提示し、顧客は、提示されたリストから選択した話者と音声合成すべき文章とをサービス提供者に伝え、サービス提供者は、顧客の選択した話者の音声特徴データに基づき、音声合成すべき文章を音声合成することを特徴とする。このようにして、サービス提供者は、インターネット等のネットワークを介して音声合成の受注を行うことができる。

ところで、ここでいう「外部」とは、ネットワークを介してサービス提供者が話者のリストを提示できる対象を示すものである。インターネットのホームページにおいては、不特定多数の者がアクセスすることが可能であり、これらのものを「外部」と称したのである。ただし、アクセスした者が必ずしも音声合成データをオーダするとは限らず、「外部」が必ずしも「顧客」なるとは限らない。

## 【 0 0 0 9 】

また、サービス提供者のサーバは、顧客に対し、生成する音声合成データについて課金し、サービス提供者は、顧客側から課金した代金を受け取った後に音声合成データを顧客に受け渡すようにしてもよい。ここでいう「顧客側」とは、顧客本人、あるいは顧客本人が契約している金融機関を指している。

## 【 0 0 1 0 】

さらに、サービス提供者は、当該サービス提供者が保有する音声特徴データの権利を有する者、つまり歌手、政治家等の有名人、テレビ番組や映画の登場人物等、特定キャラクタの音声の音源についての著作権の権利者(本人や管理者等)に対し、音声合成データの生成に応じた代金を支払うことにより、権利者に対するロイヤリティ等、権利に対する対価の支払を確実に行うことができる。

【 0 0 1 1 】

加えて、顧客は、サービス提供者から受け渡された音声合成データを機器に入力することにより、この機器において、オーダした音声合成データに基づく音声を出力することが可能となる。

【 0 0 1 2 】

これとは他に、サービス提供者は、顧客の選択に基づく音声合成データを生成するとともに、生成した音声合成データを顧客が選択した機器に入力してもよい。このようにして、サービス提供者から顧客に対し、オーダに応じた音声合成データが入力された機器を顧客に受け渡すようにしても良い。

【 0 0 1 3 】

本発明は、話者の音声进行分析して得られた音声特徴データを格納する音声特徴データ格納部と、ネットワークを介して、顧客が選択した話者および顧客が入力した文章等、顧客からの要求を受け付ける要求受付部と、顧客からの要求に応じ音声特徴データに基づいて音声合成する音声合成データ生成部と、を備えることを特徴とするサーバとして捉えることもできる。

【 0 0 1 4 】

また、音声特徴データ格納部は、音声特徴データとして、話者毎の声質データと韻律データとを格納することを特徴とする。

【 0 0 1 5 】

さらに、顧客によってなされた要求に基づき、生成する音声合成データの価格を設定する価格設定部を備えても良い。

【 0 0 1 6 】

本発明は、外部から音声合成データを生成するための要求を受け付ける処理と、前記要求に対して取引番号を生成して出力する処理と、取引番号が入力されたときに、要求に応じて生成した音声合成データを出力する処理とをコンピュータに実行させるプログラムを記憶したことを特徴とする記憶媒体として捉えることもできる。

さらに、前記プログラムは、生成した音声合成データに、音声合成データの内容を証明する認証データを添付することにより、生成された音声合成データの不

法作成や不法コピー等を防止することができる。このときに添付する認証データは、電子透かし等、いかなる形態のものであっても良い。また証明する内容としては、例えば音声合成データの出所や、音声の音源についての権利者の正式の許可を得ていること等がある。

## 【 0 0 1 7 】

また、本発明のプログラムは、音声合成すべき文章と話者の指定とを含む外部からの要求を受け付ける処理と、受け付けた要求に応じ、話者に対応した音声特徴データに基づいて文章を音声合成する処理と、をコンピュータに実行させることを特徴とすることもできる。

## 【 0 0 1 8 】

本発明は、コンピュータに実行させるプログラムが記憶された記憶手段と、記憶手段から読み出したプログラムをコンピュータに送信する送信手段とを備え、前記プログラムは、複数の音声データのリストを顧客に対して出力する処理と、顧客がリストから選択した音声特徴データに基づいて顧客が入力した文章を音声合成した音声合成データを、顧客に出力する処理をコンピュータに実行させることを特徴とするプログラム伝送装置とすることもできる。

## 【 0 0 1 9 】

また、本発明を、サービス提供者が顧客の指定に応じて生成した音声合成データが格納されていることを特徴とする音声合成データ記憶媒体として捉えることもできる。ここで、音声合成データ記憶媒体としては、フレキシブルディスク、CD-ROM、DVD、メモリ、ハードディスク等の各種記憶媒体がある。そして、このような音声合成データ記憶媒体に格納した音声合成データは、コンピュータや携帯型電話端末、カーナビゲーション装置等の様々な機器に転送し、この機器で前記音声合成データに基づく音声を出力すれば良い。また、音声合成データ記憶媒体として、可搬性のあるメモリを用いれば、サービス提供者から顧客に対して音声合成データを受け渡す際にも本発明を適用できる。

## 【 0 0 2 0 】

さらに、本発明は、指定された話者および文章に基づいてサービス提供者が生成した音声合成データを記憶する記憶部と、当該記憶部で記憶した音声合成デー

タに基づく音声を出力する音声出力部と、を備えることを特徴とする音声出力機器とすることもできる。このような音声出力機器としては、生成した音声合成データを組み込んだ(入力した)、玩具、目覚し時計、携帯型電話端末、カーナビゲーション装置、メモリープレーヤ等の音声再生機器等がある。

#### 【 0 0 2 1 】

##### 【発明の実施の形態】

以下、添付図面に示す実施の形態に基づいてこの発明を詳細に説明する。

図 1 は、本実施の形態におけるシステム構成を説明するための図である。この図において、符号 1 は、音声合成データを提供するサービス提供者であり、本実施の形態におけるシステムのウェブサーバとして機能している。符号 2 は歌手、政治家等の有名人、テレビ番組や映画の登場人物等の、特定キャラクタの音源の権利(著作権等)を所有あるいは管理する権利者であり、サービス提供者 1 との間で、音声データの使用許諾、および音声データの使用に応じたロイヤリティの支払条件等に関し、所定の契約を結んでいる。符号 3 は音声合成データを購入したいと欲する購入者である顧客(外部、顧客側)、符号 4 はサービス提供者 1 と提携している金融機関(顧客側)であり、例えばクレジットカード会社やデビットカード(Debit Card)等の即時決済サービスを行っている銀行等である。また符号 5 はインターネット等のネットワークであり、このネットワーク 5 には、ウェブサーバとしてのサービス提供者 1、ウェブ端末である顧客 3 がそれぞれ接続されている。

#### 【 0 0 2 2 】

顧客 3 のウェブ端末は、例えば、ウェブブラウザ等のソフトウェアを備えた PC であり、サービス提供者 1 のホームページの閲覧や、送られてくる各種情報等をディスプレイ上で表示することができ、また送られてくる画面において、各種情報や金額等を入力できるポインティングデバイスやキーボード等の入力手段を備えている。

#### 【 0 0 2 3 】

また、金融機関 4 は、サービス提供者 1 と、ネットワーク 5 あるいは他のネットワーク等を介して接続されており、双方間で情報の送受を行うことができるよ

うになっている。またこの金融機関 4 は、顧客 3 との間で所定の契約を結んでいるものとする。

【 0 0 2 4 】

さて、本実施の形態におけるサービス提供者 1 は、顧客 3 のオーダーに応じ、顧客 3 の指定した特定キャラクタ(以下、スピーカと称する)の音声で、顧客 3 の希望する文章を出力(発声)させるための音声合成データを提供する。

【 0 0 2 5 】

図 2 は、ウェブサーバであるサービス提供者 1 のサーバ構成を示すブロック図である。図 2 において、ネットワーク 5 に対する送受部として例えば用いられる HTTP サーバ 11 は、ネットワーク 5 を介して外部のウェブ端末とデータの送受信を行っている。このサーバは、大きく分けると、顧客情報に関する処理を行う顧客情報処理ブロック 20、顧客 3 からのオーダー、支払、顧客 3 に対してのデリバリ(受け渡し)等の処理を行うオーダー/支払/デリバリ処理ブロック 30、権利者 2 との間でのロイヤリティ契約に基づく処理を行うロイヤリティ処理ブロック 40、音声合成データ生成を行うためのコンテンツを処理するコンテンツ処理ブロック 50、顧客 3 からのオーダーを受けて音声合成データを生成する音声合成データ生成ブロック 60 から構成されている。また、このサーバは、顧客 3 との取引によって生じる金銭の授受処理を行うため、支払ゲートウェイ 70、ロイヤリティゲートウェイ 75 を備えており、この支払ゲートウェイ 70、ロイヤリティゲートウェイ 75 を介して、このサーバは、サービス提供者 1 がサーバ外に設けたロイヤリティ支払システム 80 とクレジットカードシステム 90 等とに接続されている。

また、このサーバは、画面データ生成部 13 を備えており、この画面データ生成部 13 は、顧客 3 が入力したデータを受け、データの種別に応じて各データをサーバ各部に割り振る。またサーバ各部から受領したデータ等を基に、画面データを生成することができる。

【 0 0 2 6 】

顧客情報処理ブロック 20 は、顧客情報処理(Customer Management)部 21、顧客情報 DB(データベース) 22 を備えており、顧客情報処理部 21 は、顧客 3

から得た、顧客名、住所、eメールアドレス等の情報を顧客情報DB22に格納し、必要に応じて格納した情報を顧客情報DB22から取り出す。

【0027】

オーダー／支払／デリバリ処理ブロック30は、オーダー処理部(要求受付部)31、支払処理部(価格設定部)32、デリバリ処理部33、オーダー／支払／デリバリDB34、デリバリサーバ35を備える。

【0028】

オーダー処理部31は、顧客3からのオーダー内容をオーダー／支払／デリバリDB34に格納するとともに、そのオーダー内容に基づく音声合成データを生成するための指令をコンテンツ処理ブロック50に出す。

【0029】

支払処理部32は、オーダー／支払／デリバリDB34に予め格納されている価格データに基づき、顧客3からのオーダー内容に応じた価格を算出して出力し、また顧客3から得たクレジットカード情報等、支払に関する情報をオーダー／支払／デリバリDB34に格納する。またこの支払処理部32では、支払ゲートウェイ70およびサーバ外部のクレジットカードシステム90を介して、金融機関4との間で、顧客3から得たクレジットカード情報の照合、金融機関4に対する代金支払請求、金融機関4からの入金確認処理等を行う。

【0030】

デリバリ処理部33は、顧客3から受けたオーダーによって生成する音声合成データがデリバリ可能となるまでのスケジュールの管理および出力、顧客3がオーダーした音声合成データを受け取るためのURL(Uniform Resource Locators)等の出力、顧客3からのオーダーに対しての取引番号の生成および出力等を行う。そして、デリバリ処理部33で顧客3に対して出力した情報を、必要に応じてオーダー／支払／デリバリDB34に格納する。

【0031】

ロイヤリティ処理ブロック40は、ロイヤリティ処理部41とロイヤリティ契約DB42とを備える。ロイヤリティ契約DB42には、権利者2との間でのロイヤリティ契約等のデータが格納されており、ロイヤリティ処理部41には、こ

のデータに基づき、顧客 3 からのオーダーに応じたロイヤリティの計算、ロイヤリティゲートウェイ 7 5 およびロイヤリティ支払システム 8 0 を介しての権利者 2 に対するロイヤリティの支払処理を行う。

【 0 0 3 2 】

コンテンツ処理ブロック 5 0 は、コンテンツ処理部(音声合成データ生成部) 5 1 とコンテンツ DB 5 2 とから構成され、コンテンツ処理部 5 1 は、オーダー処理部 3 1 から転送されたオーダー内容情報、指定されたスピーカ、指定された文章等をコンテンツ DB 5 2 に格納するとともに、後述する音声合成データ生成ブロック 6 0 で生成された音声合成データを出力する処理を行う。

また、コンテンツ DB 5 2 には、登録されているスピーカ(音声)のリスト、およびその一部ないし全部の音声サンプルデータが格納されており、コンテンツ処理部 5 1 では、顧客 3 からの要求に応じ、指定された音声サンプルデータを出力できるようになっている。

【 0 0 3 3 】

音声合成データ生成ブロック 6 0 は、音声合成処理部(音声合成データ生成部) 6 1、音声特徴 DB (音声特徴データ格納部) 6 2 を備えて構成される。

音声特徴 DB 6 2 には、予め登録されたスピーカの音声データ(音声特徴データ)が格納されている。この音声データは、登録されたスピーカ一人一人について、そのスピーカの声の声質をデータ化した声質データ D 1 と、そのスピーカが話すときの韻律をデータ化した韻律データ D 2 からなり、これら声質データ D 1、韻律データ D 2 が、登録されたスピーカ毎に分類されて音声特徴 DB 6 2 に格納されているのである。

【 0 0 3 4 】

図 3 に示すように、音声特徴 DB 6 2 に格納されている音声データを得るには、まず、スピーカが直接録音したもの、あるいはテレビ番組や映画等を録画したもの、歌等を録音したもの等をデータ化し、音源データとして保持する。そしてその音源データを解析し、そのスピーカの音声の特徴、つまりスピーカの音声の声質とスピーカが話すときの韻律とを抽出し、それぞれ声質データ D 1、韻律データ D 2 としてデータ化するのである。

## 【 0 0 3 5 】

図 2 に示したように、音声合成処理部 6 1 は、文章の分析処理を行うテキストマイニングエンジン 6 3、音声合成データを生成するシンセサイズエンジン 6 4、生成した音声合成データに電子透かし等の認証 (Water Mark) を埋め込むウォーターマークエンジン 6 5、音声合成データをファイル形式に変換するファイルフォーマットエンジン 6 6 を備えて構成されている。

## 【 0 0 3 6 】

このような音声合成処理部 6 1 において、音声合成データを生成するには、コンテンツ DB 5 2 から、顧客 3 からのオーダーにおいて指定されたスピーカを示すデータを引き出す。また、音声特徴 DB 6 2 からは、このスピーカの音声データ (声質データ D 1 と韻律データ D 2) を引き出し、さらに、コンテンツ DB 5 2 から、顧客 3 からのオーダーにおいて指定された文章を引き出す。

そしてまず、図 3 に示したように、顧客 3 の入力した文章を、テキストマイニングエンジン 6 3 に備えた文法 DB 6 7 に格納された文法等に基づいて分析する (ステップ S 1)。次いで、シンセサイズエンジン 6 4 で、その分析結果と韻律データ D 2 とに基づいて、入力された文章に応じた韻律制御を行い (ステップ S 2)、スピーカの韻律を反映させる。続いて、スピーカの韻律を反映させたデータに、スピーカの声質データ D 1 を組み合わせて生成した音声波形により、所定の音声合成データを得る (ステップ S 3)。ここで、所定の音声合成データとは、言うまでもなく、顧客 3 からのオーダーによって指定されたスピーカの音声で、指定された文章内容を出力 (発声) するための音声データである。

## 【 0 0 3 7 】

次に、ウォーターマークエンジン 6 5 において、生成された音声合成データが、正式に認証されているもの、つまり音源権利者の許諾を得ていることを証明する電子透かし (認証データ) を、音声合成データに埋め込む (ステップ S 4)。

しかる後、ファイルフォーマットエンジン 6 6 で、生成された音声合成データを、所定のファイル形式、例えば W A V 形式等のサウンドファイルに変換し、さらに、この音声合成データが、顧客 3 からオーダーされたものであることを示すファイル名等を付与する。



## 【0038】

このようにして音声合成処理部61で生成された音声合成データは、音声合成処理部61から出力され(ステップS5)、顧客3がこれをダウンロードするまで、コンテンツDB52に格納される。このときには、生成された音声合成データと、顧客3がオーダー時に付与された取引番号とが、関連付けされてコンテンツDB52に格納される。

## 【0039】

ところで、上記のように、実際の音声から声質データD1、韻律データD2を抽出する技術、声質データD1、韻律データD2から希望の文章を話す音声合成データを生成する技術、としては、既に様々なものが提案あるいは実用化されているので、それらの技術を適宜用いれば良く、ここではより具体的に限定する意図はない。ここで、このような技術の一例を挙げれば、例えば特開平9-90970号公報に開示された技術があり、ここに開示された技術を適用すれば、上記の如く特定スピーカの音声合成を行うことができる。しかし、前記公報に開示された技術は、あくまでもその一例に過ぎず、適宜他の技術を用いることも可能である。

## 【0040】

さて、上記したようなシステムにおいて、顧客3が希望の音声合成データを購入する方法を、図4～図6を参照しつつ説明する。

## 〔オーダーセッション〕

図4では、サービス提供者1と顧客3との間でのやり取りを示している。この図4に示すように、まず、顧客3は、ウェブ端末からインターネット等のネットワーク5を介し、サービス提供者1のウェブサーバにアクセスする(ステップS11)。するとサービス提供者1のオーダー処理部31側から顧客3に対して、スピーカ選択要求がなされる(ステップS21)。このときには、顧客3のウェブ端末のディスプレイ上に、サービス提供者1のコンテンツDB52に登録されているスピーカのリストが表示される。このリストには、スピーカの名称が、ジャンル別、アルファベット順、あいうえお順等で表示され、さらに各スピーカの名称とともに、スピーカの顔の画像やアニメーションを表示しても良い。顧客3はこ

のリストから希望のスピーカ(特定の話者、音源)を選択し、ディスプレイ上に表示されたボタン等でこれを入力する(ステップS 1 2)。また、スピーカの選択にあたっては、顧客3が選択したスピーカの音声サンプルデータを、コンテンツDB 5 2からダウンロードすることもできる。顧客3は、必要に応じ音声サンプルデータをダウンロードして音声サンプルを試聴し、最終的に選択するスピーカを決定する。

## 【0 0 4 1】

スピーカの選択が行われると、サービス提供者1のオーダ処理部3 1から顧客3に対し、文章入力要求が行われる(ステップS 2 2)。顧客3は、ディスプレイ上に表示される入力欄に、希望する文章をキーボード等の入力手段で入力する(ステップS 1 3)。

## 【0 0 4 2】

すると、サービス提供者1のオーダ処理部3 1側では、テキストマイニングエンジン6 3において入力された文章を分析し、文章のリーガルチェックを行うとともに、文章を構成する文字数あるいは単語数をカウントする。さらに、ロイヤリティ契約DB 4 2に、ステップS 1 2で選択されたスピーカのロイヤリティを含んだベース価格を参照する。そして、支払処理部3 2が、前記文字数あるいは単語数のカウント値と、選択されたスピーカとに応じたベース価格とから、顧客3のオーダ内容に応じた価格を算出する。

しかる後、オーダ処理部3 1では、顧客3からのオーダ内容、つまり選択したスピーカの名称、入力した文章を表示するとともに、オーダ内容に応じた価格を表示し、顧客3に対してオーダ内容の確認を要求する(ステップS 2 3)。顧客3は、サービス提供者1側から表示されたオーダ内容を確認し、ディスプレイ上に表示されたボタン等で、確認の意思表示を入力する(ステップS 1 4)。

## 【0 0 4 3】

次に、サービス提供者1のオーダ処理部3 1は、顧客3に対し、顧客情報の入力を要求する(ステップS 2 4)。顧客3は、自分の氏名、住所、eメールアドレス等を必要に応じて入力する(ステップS 1 5)。するとサービス提供者1側では、顧客情報処理部2 1で、入力された顧客3の情報を顧客情報DB 2 2に格納

する。

【0044】

続いて、サービス提供者1のオーダー処理部31から顧客3に対し、支払情報の入力及要求されるので(ステップS25)、顧客3は、クレジットカードの種類、クレジットカード番号等を入力する(ステップS16)。このとき、例えばデビットカード等の即時決済システムが利用可能であれば、支払情報として、銀行のキャッシュカードの番号と、暗証番号とを入力するようにしても良い。

なお、上記ステップS15、S16において、顧客3が予めサービス提供者1にメンバー登録をしているのであれば、ステップS11のアクセス(ログイン)時、あるいはステップS16において、メンバーIDやパスワードを入力することにより、上記ステップS16の顧客情報入力およびステップS17の支払情報の入力を省略することも可能である。

【0045】

顧客3から支払情報を受けたサービス提供者1側は、支払処理部32において、支払ゲートウェイ70およびクレジットカードシステム90を介し、金融機関4に対して顧客3の支払情報を照会する(ステップS26)。これを受けた金融機関4側では、顧客3の支払情報を確認し、サービス提供者1側に対し確認(承認)の可否を回答する(ステップS30)。金融機関4から確認が得られた後、支払処理部32では顧客3の支払情報をオーダー/支払/デリバリDB34に格納する。

【0046】

次いで、サービス提供者1のオーダー処理部31から顧客3に対し、オーダーの最終確認を要求し(ステップS27)、顧客3は、オーダーの最終確認を行う(ステップS17)。

【0047】

顧客3からの最終確認が得られると、サービス提供者1のオーダー処理部31においてオーダーを受け付け(ステップS28)、そのオーダー内容をコンテンツ処理部51に送る。これとともに、デリバリ処理部33で、顧客3からの当該オーダーに対する取引番号(Transaction ID)を生成する。この取引番号は、顧客3からのオーダー一つ一つに対して個別に生成されるものである。そして、オーダー処理部31

では、デリバリ処理部 3 3 で生成された取引番号とともに、後に顧客 3 が音声合成データをダウンロードするためのサイトの URL、音声合成データが完成してデリバリ可能となるスケジュール(完成予定期日)を顧客 3 に対し出力する(ステップ S 2 9)。また、HTTP サーバ 1 1 から、顧客 3 に対し、完成した音声合成データのダウンロード方法を転送する。顧客 3 は、これを受け取ることにより、オーダセッションが終了する。

#### 【 0 0 4 8 】

上記の如くして顧客 3 からのオーダを受けたサービス提供者 1 側では、オーダー内容に基づき、前記したような手順で音声合成データを生成する。その一方、サービス提供者 1 側では、金融機関 4 に対し、顧客 3 のオーダーに応じた代金の決済を要求する。その要求は、顧客 3 からのオーダーを受け付けた後であれば、音声合成データの生成前、生成中、生成後、音声合成データの顧客 3 へのデリバリ後、のいずれであっても良いが、一例を挙げれば図 5 に示すような処理となる。

この図 5 に示すように、サービス提供者 1 側では、顧客 3 とのオーダセッションの終了後、支払処理部 3 2 において、支払ゲートウェイ 7 0 およびクレジットカードシステム 9 0 を介して、金融機関 4 に対し顧客 3 のオーダーに応じた代金の決済を要求する(ステップ S 4 1)。金融機関 4 では、この要求に応じ、サービス提供者 1 に対し代金の支払を行う(ステップ S 5 0)。サービス提供者 1 側において、金融機関 4 からの代金支払を確認した時点で、音声合成データの作成に着手する(ステップ S 4 2)。そして、音声合成データの生成が完了した時点で、生成した音声合成データを、コンテンツ DB 5 2 に格納(保管)する(ステップ S 4 3)。

#### 【 0 0 4 9 】

##### 〔ダウンロードセッション〕

前記オーダーセッションのステップ S 2 9 においてサービス提供者 1 側から告知された完成予定期日以降、顧客 3 がオーダーした音声合成データを受け取るには、図 6 に示す流れを経る。

この図 6 に示すように、顧客 3 が、サービス提供者 1 のサーバの、オーダーセッションのステップ S 2 9 で示された URL にアクセスすると(ステップ S 6 1)、

サービス提供者 1 のコンテンツ処理部 5 1 では顧客 3 に対して取引番号の入力を要求する(ステップ S 7 1)。顧客 3 は、オーダセッションのステップ S 2 9 でサービス提供者 1 から示された取引番号を入力する(ステップ S 6 2)。この取引番号が、オーダした音声合成データを受け取るためのいわば合鍵となるので、取引番号が一致しない限り、音声合成のデータの受け取りは不可能である。

## 【0050】

顧客 3 が入力した取引番号が、オーダ／支払／デリバリ DB 3 4 に格納されている取引番号と一致すると、デリバリ処理部 3 3 は、顧客 3 に対し、オーダ／支払／デリバリ DB 3 4 に格納されている顧客 3 のオーダ内容を表示する。表示するオーダ内容としては、顧客 3 の名前、オーダしたスピーカの名称、オーダした文章等の情報がある。これとともに、デリバリ処理部 3 3 では、顧客 3 側のディスプレイ上に、オーダされた音声合成データのファイルをダウンロードするためのボタン等を表示し、顧客 3 からのダウンロード開始信号の入力を要求する(ステップ S 7 2)。

そして顧客 3 が、ディスプレイ上のボタンを操作すると、サービス提供者 1 側に対して、オーダした音声合成データのファイルのダウンロードを開始する信号が入力される(ステップ S 6 3)。

## 【0051】

すると、サービス提供者 1 側では、コンテンツ処理部 5 1 において、コンテンツ DB 5 2 に所定のファイル形式で格納されている、顧客 3 のオーダに応じて生成した音声合成データのファイルを、顧客 3 に対して出力する(ステップ S 7 3)。顧客 3 側では、出力された音声合成データのファイルをダウンロードする(ステップ S 6 4)。ダウンロードが完了した時点で、音声合成データのダウンロードセッションが完了し、つまり顧客 3 のオーダに対するサービス提供者 1 との間での取引がすべて完了する。

## 【0052】

一方、上記オーダセッションとは別に、金融機関 4 は、顧客 3 に対し、代金の支払を要求し、顧客 3 はこれに応じて金融機関 4 に代金の支払を行う。

またサービス提供者 1 は、顧客 3 のオーダ内容に応じたロイヤリティを、権利

者 2 に対し別途支払う。

【0053】

しかる後は、顧客 3 は、自らの PC 端末において、ダウンロードした音声合成データのファイルを保存し、これを再生用のソフトウェアによって再生しても良い。また、図 1 に示したように、顧客 3 が、音声合成データを記憶する記憶部と、当該記憶部で記憶した音声合成データに基づく音声を出力可能な音声出力部とを備えた機器(音声出力機器) 100、例えば玩具、目覚し時計、携帯型電話端末、カーナビゲーション装置、いわゆるメモリープレーヤのような音声データ再生機器等、を別途購入、あるいは既に所有している場合、ダウンロードした音声合成データを機器 100 に転送し、この機器 100 において、音声合成データを再生するようにしても良い。このとき、各種の機器 100 に音声合成データを転送するには、例えばデータ伝送用の接続ケーブルや無線、赤外線通信を用いても良いし、また可搬性のあるメモリ(音声合成データ記憶媒体)に音声合成データを格納し、このメモリを介して機器 100 に音声合成データを転送してもよい。

なお、図 1 は、上記したような音声合成データのオーダー～受け取りまでの流れを示すものであり、この図 1 中、①～⑥は音声合成データの提供までに着目した処理の順序を示している。

【0054】

上述したようにして、顧客 3 は、オーダーした音声合成データにより、歌手、政治家等の有名人、テレビ番組や映画の登場人物等、希望したスピーカの音声で、希望した文章を、自ら所有する PC や機器 100 で出力させることが可能となる。つまり、例えば目覚し時計の目覚しアラーム、携帯型電話端末の呼び出し音(メッセージ)や留守応答メッセージ、カーナビゲーション装置におけるガイダンスメッセージ等、を自分の好みに応じて変更することが可能となるのである。

【0055】

しかも、音声合成データは、顧客 3 のオーダーに応じて生成され、取引番号によって顧客 3 に対して受け渡されるので、唯一無二なものとなる。その際、顧客 3 のオーダーに応じた価格が設定され、音源の権利者 2 に対してのロイヤリティの支払を確実に行うことができる。

## 【 0 0 5 6 】

さらに、顧客 3 によって音声合成データを入力できる機器 1 0 0 に対し、上記システムにより、顧客 3 が機器 1 0 0 で再生させるメッセージ等を適宜変更することが可能となる。つまり、顧客 3 が新たに音声合成データをオーダしてこれを入力した後、機器 1 0 0 に入力されている音声合成データを、新たな音声合成データに入れ換えるのである。このようにして、上記システムにより、機器 1 0 0 に対する顧客 3 の飽きを防ぎ、機器 1 0 0 の付加価値をさらに高めることが可能となる。

## 【 0 0 5 7 】

なお、上記実施の形態において、例えばデリバリ処理部 3 3 で、完成予定期日を顧客 3 に告知し、顧客 3 は完成予定期日以降に音声合成データを受け取る構成となっているが、顧客 3 からオーダを受けたセッションの間(例えばオーダを受けて即時)に、音声合成データを提供できるのであれば、そのような処理は不要である。

また、図 4 ～図 6 に示した処理の流れにおいて、所定の入力や確認が行われないうときには、処理が中止、あるいは前のステップに戻る処理が行われたりするの言うまでもない。

## 【 0 0 5 8 】

## 〔他の実施の形態〕

次に、図 7 を参照しつつ他の実施の形態について説明する。以下の説明において、上記実施の形態と同様の構成については、同符号を付してその説明を省略する。

この図 7 に示す実施の形態では、サービス提供者 1 が顧客 3 に対し、音声合成データのみではなく、顧客 3 がオーダした音声合成データを入力した機器を提供するのである。この図 7 は、上記したような音声合成データのオーダ～受け取りまでの流れを示すものであり、図中、①～⑤は音声合成データの提供までに着目した処理の順序を示している。

## 【 0 0 5 9 】

これには、サービス提供者 1 は、顧客 3 に対し、スピーカのリストと機器のリ

ストとを提供する。ここで、顧客 3 がオーダーする機器としては、玩具、目覚し時計、ナビゲーション装置等、音声合成データを入力する機器であればいかなるものであっても良い。

顧客 3 は、サービス提供者 1 に対し、音声合成データのオーダーを上記実施の形態と同様に行うとともに、音声合成データを入力する機器のオーダーを行う。この機器のオーダーは、上記実施の形態で示したオーダーセッション(図 4 参照)の中の適宜タイミングで行えば良い。そして、顧客 3 に対しては、サービス提供者 1 から、オーダーした音声合成データと、選択した機器とに応じた価格が提示されるので、顧客 3 がそのオーダー内容を確認し、その意思表示をサービス提供者 1 に対して示した時点で、オーダーが完了する。

【 0 0 6 0 】

そして、サービス提供者 1 は、顧客 3 のオーダーに応じ、上記実施の形態の如く音声合成データを生成するとともに、生成された音声合成データを、顧客 3 がオーダーした機器に入力し、この機器を顧客 3 に納品する。またサービス提供者 1 は、顧客 3 が指定した金融機関 4 に対し、顧客 3 がオーダーした音声合成データと機器の代金の支払を請求し、代金の決済を行う。

また顧客 3 は、金融機関 4 に対し、オーダー内容に応じた代金を支払い、サービス提供者 1 が音声合成データの生成に応じたロイヤリティを権利者 2 に支払うことにより、すべての取引が完了する。

【 0 0 6 1 】

なお、上記各実施の形態において、サービス提供者 1 と金融機関 4 との間、金融機関 4 と顧客 3 との間での決済タイミングについては、何ら問うものではなく、上記した記載に限定する意図はない。また、顧客 3 によるサービス提供者 1 への代金支払は、必ずしも金融機関 4 を介さなくても良く、例えば電子マネーやプリペイドカード等によるものでも良い。

【 0 0 6 2 】

また、上記実施の形態で示したように、顧客 3 は、音声合成データのみを購入しても良いし、音声合成データを組み込んだ機器 1 0 0 を購入してもよいが、さらに、顧客 3 が購入した音声合成データを機器メーカーに送り、当該機器メーカーで



前記音声合成データを顧客 3 が希望する機器に組み込み、これを顧客 3 に販売する形態、顧客 3 のオーダにより生成した音声合成データを、サービス提供者 1 が機器メーカーに送り、当該機器メーカーで前記音声合成データを機器に組み込み、これを顧客 3 に受け渡す形態も可能である。

【 0 0 6 3 】

また、音声合成データは、音声のみによるメッセージに限らず、歌(伴奏有り／なし)や朗読等であっても良い。さらに、顧客 3 は、文章内容を自由にオーダすることも可能であるが、例えばサービス提供者 1 側が提示した文章の中から顧客 3 が選択する構成としても良い。このような構成においては、文章として例えば詩や小説等をサービス提供者 1 が提示し、顧客 3 がスピーカを選択することにより、いわば顧客 3 が、好きなスピーカによる朗読の音声合成データを得ることのできるシステム構成とすることも可能である。

【 0 0 6 4 】

またサービス提供者 1 から顧客 3 に対して提供する音声合成データは、上記実施の形態で示したように、オンラインでの伝送(ダウンロード)、機器に入力した状態での受け渡しに限らず、例えばフレキシブルディスク等、種々の形態の記憶媒体(音声合成データ記憶媒体)に格納して受け渡すことも可能である。

【 0 0 6 5 】

加えて、上記のようなプログラムをコンピュータに実行させるため、前記プログラムを記憶した、CD-ROM、DVD、メモリ、ハードディスク等の記憶媒体の形態をなしていても良い。また上記のようなプログラムを記憶するCD-ROM、DVD、メモリ、ハードディスク等の記憶手段と、この記憶手段から当該プログラムを読み出し、当該プログラムを実行する装置側に直接あるいは間接的に当該プログラムを送信する送信手段とを備える伝送装置の形態を成していても良い。

【 0 0 6 6 】

【発明の効果】

以上説明したように、本発明によれば、顧客は、希望したスピーカの音声で、希望した文章内容の音声合成データを得ることができ、また音源の権利者に対し

てのロイヤリティの支払を確実に行うことができる。

【図面の簡単な説明】

【図 1】 本実施の形態におけるシステム構成を示す図である。

【図 2】 サービス提供者のサーバ構成を示すための図である。

【図 3】 サービス提供者における音声合成データ生成方法を示すための図である。

【図 4】 音声合成データを顧客がオーダーするときの処理の流れを示す図である。

【図 5】 音声合成データを生成するときの処理の流れを示す図である。

【図 6】 オーダーした音声合成データを顧客に受け渡すときの処理の流れを示す図である。

【図 7】 他の実施の形態におけるシステム構成を示す図である。

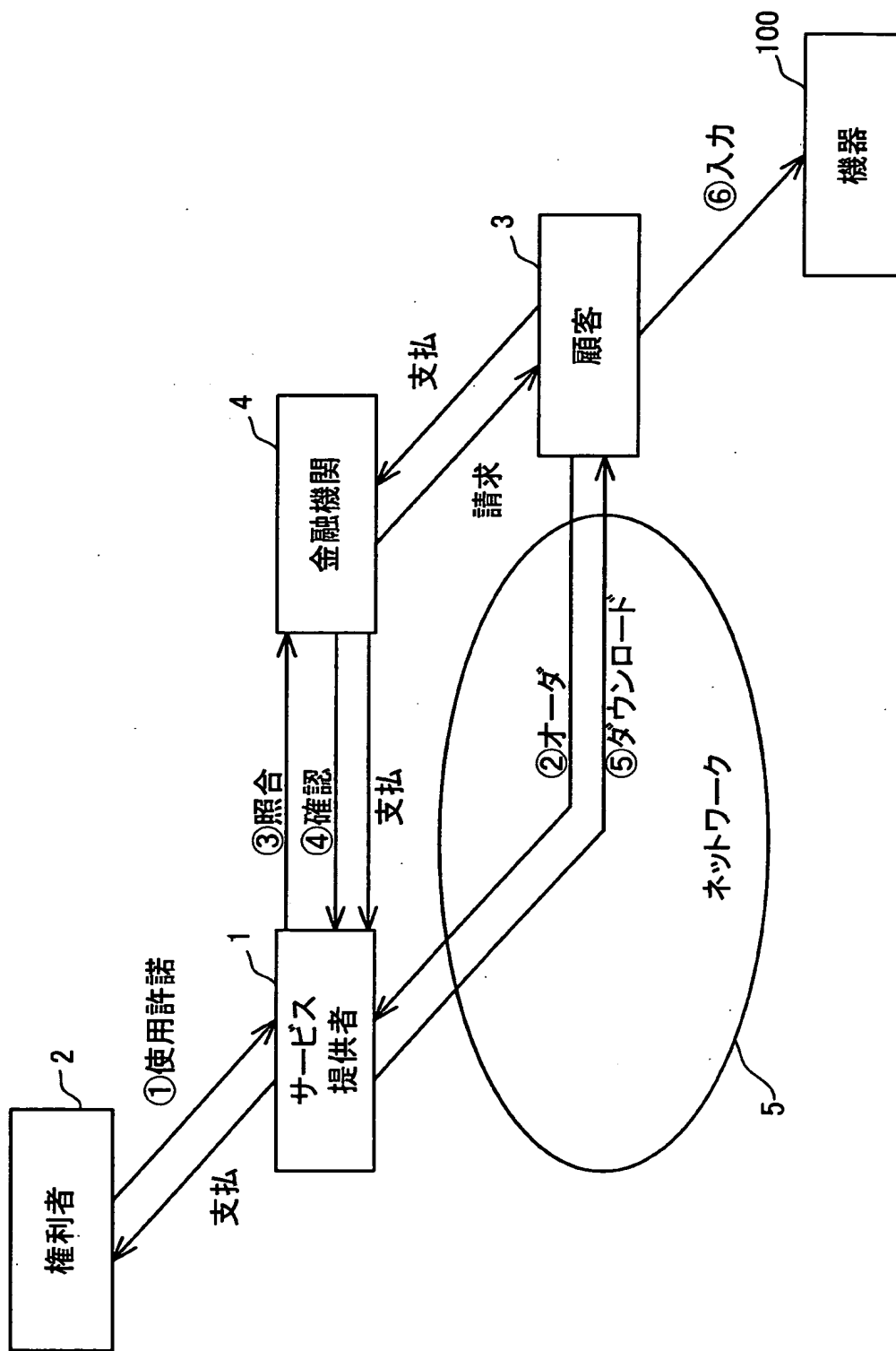
【符号の説明】

1 …サービス提供者、2 …権利者、3 …顧客(外部、顧客側)、4 …金融機関(顧客側)、5 …ネットワーク、2 1 …顧客情報処理部、2 2 …顧客情報DB、3 1 …オーダー処理部(要求受付部)、3 2 …支払処理部(価格設定部)、3 3 …デリバリ処理部、3 4 …オーダー/支払/デリバリDB、4 1 …ロイヤリティ処理部、4 2 …ロイヤリティ契約DB、5 1 …コンテンツ処理部(音声合成データ生成部)、5 2 …コンテンツDB、6 1 …音声合成処理部(音声合成データ生成部)、6 2 …音声特徴DB(音声特徴データ格納部)、8 0 …ロイヤリティ支払システム、9 0 …クレジットカードシステム、1 0 0 …機器(音声出力機器)、D 1 …声質データ、D 2 …韻律データ

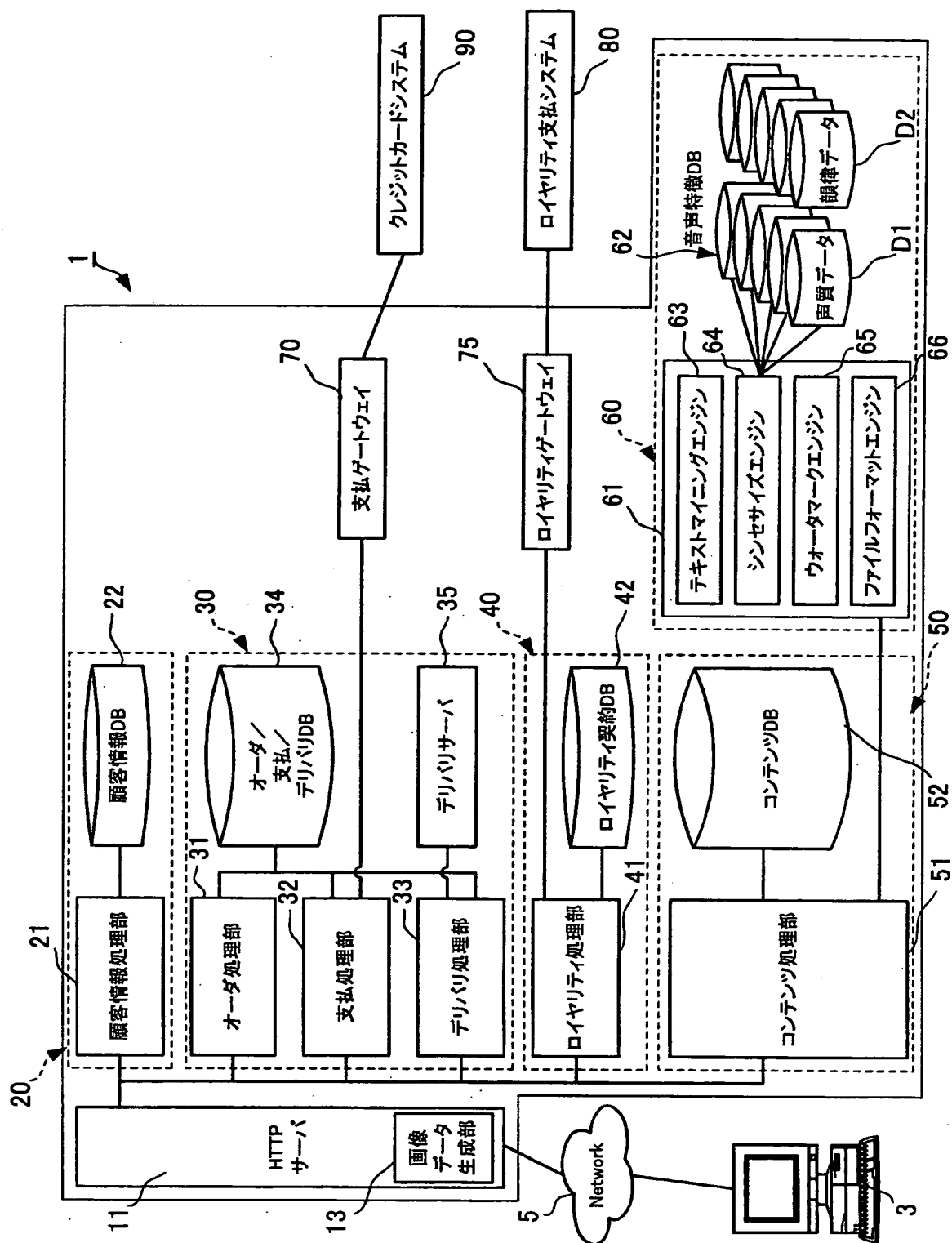
【書類名】

図面

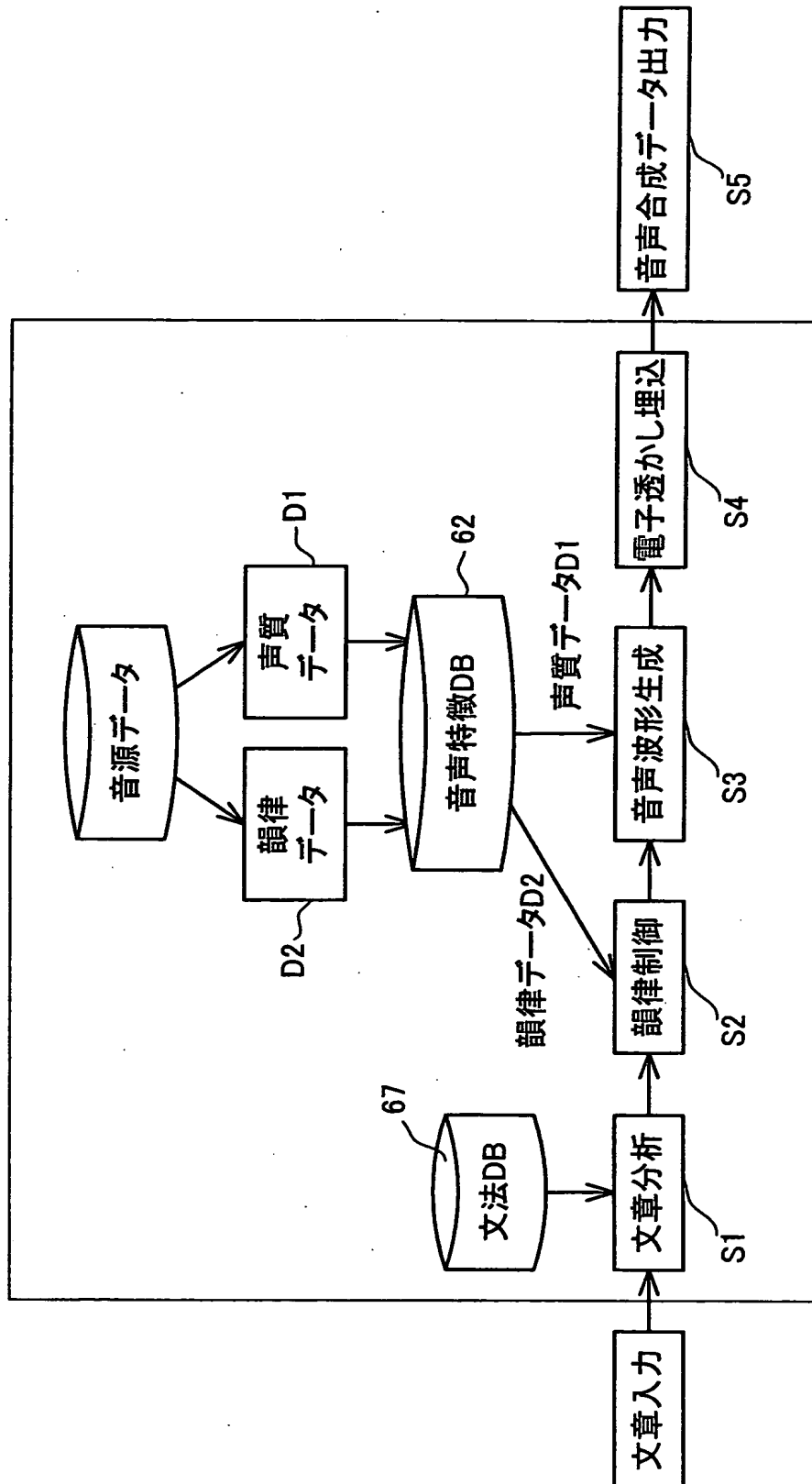
【図 1】



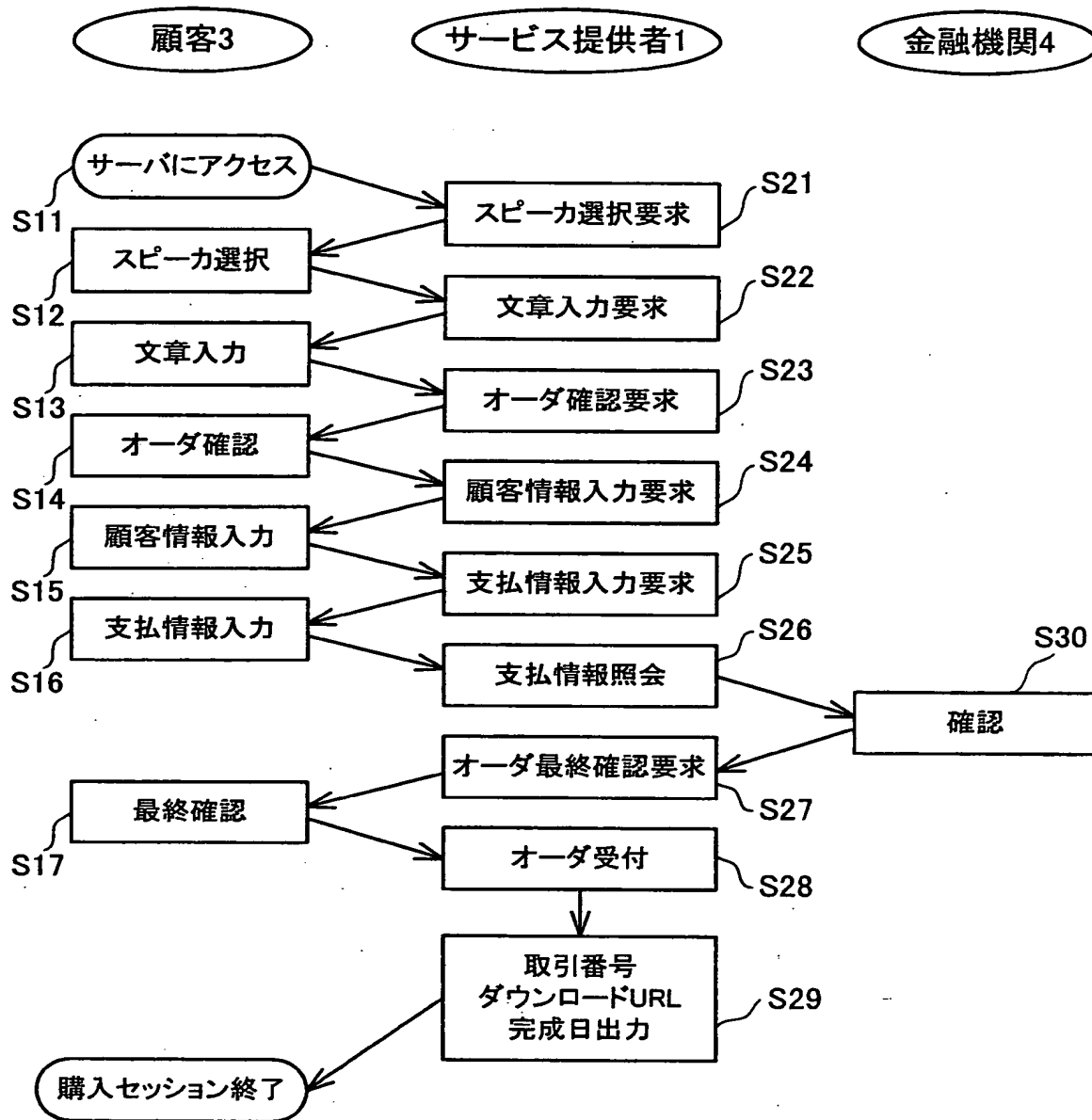
【図 2】



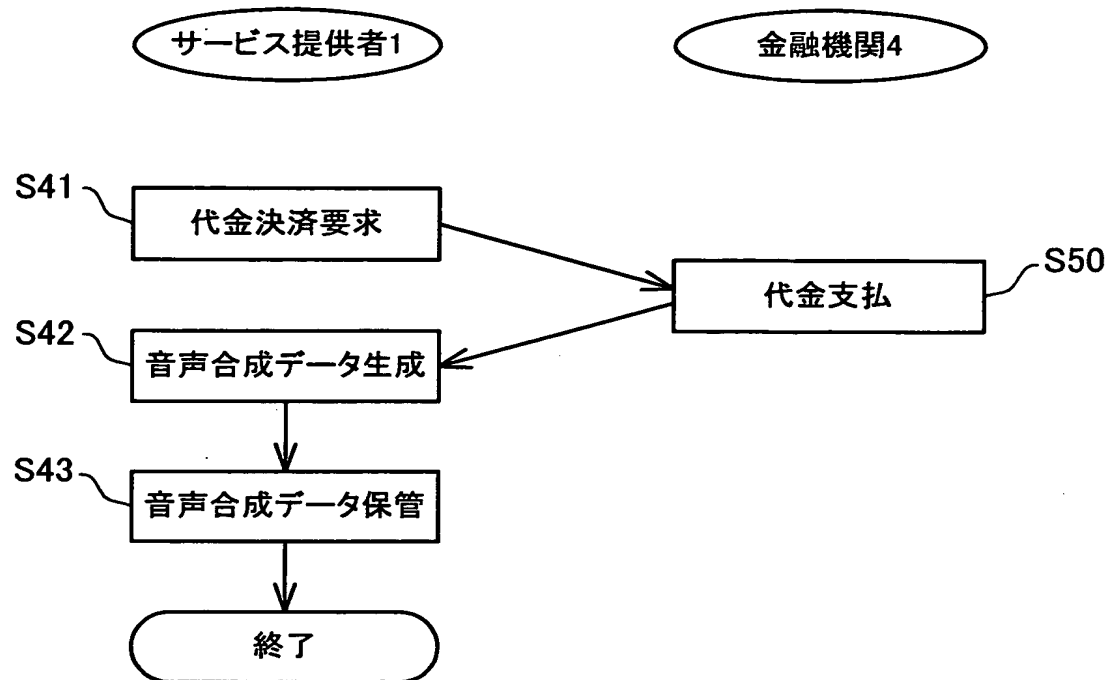
【図 3】



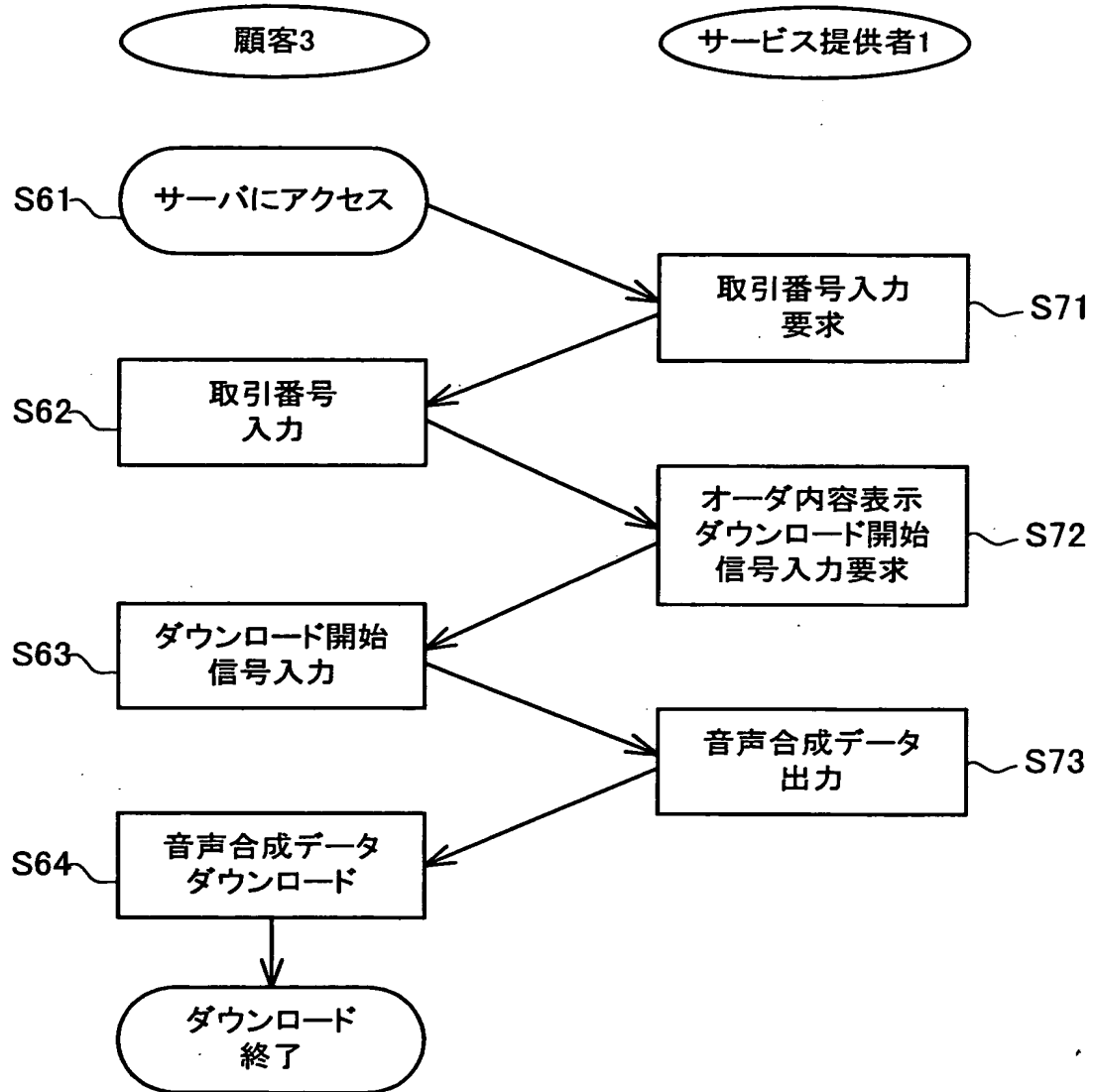
【図 4】



【図 5】

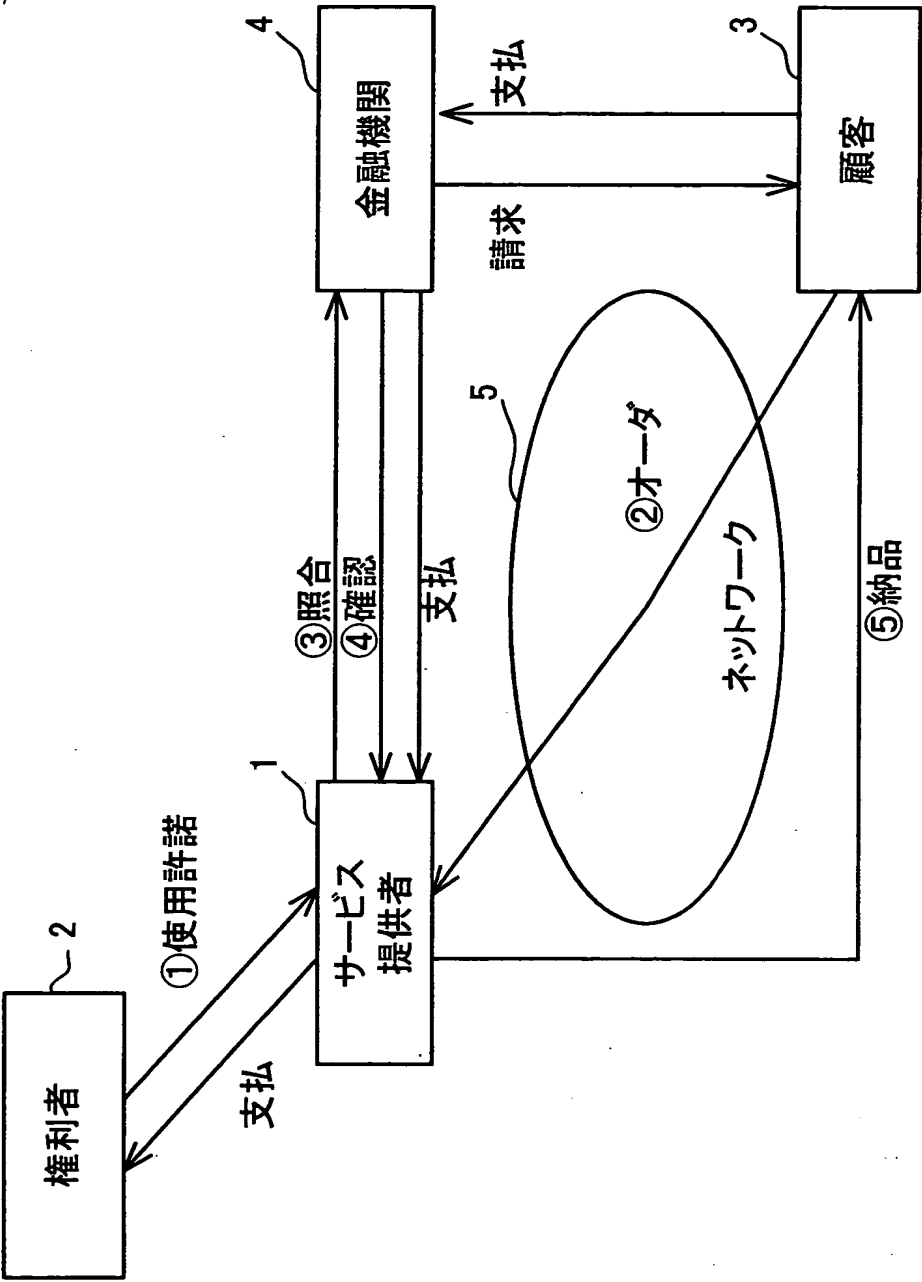


【図 6】





【図 7】



【書類名】 要約書

【要約】

【課題】 顧客の好みに応じた音声合成メッセージを提供する音声合成システム、音声合成方法、サーバ、記憶媒体、プログラム伝送装置、音声合成データ記憶媒体、音声出力機器を提供する。

【解決手段】 サービス提供者 1 は、ネットワーク 5 を介してなされた顧客 3 のオーダーに応じ、顧客 3 が選択したスピーカの音声特徴データと、顧客 3 が入力した文章とに基づく音声合成データを生成した後、顧客 3 に受け渡す構成とした。このとき、顧客 3 のオーダーに対して取引番号が付与され、取引番号が顧客 3 から提示されたときに、生成した音声合成データを顧客 3 に受け渡すようにする。顧客 3 は、受け取った音声合成データを機器 1 0 0 に入力し、オーダーした音声を出力する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [390009531]

1. 変更年月日 2000年 5月16日

[変更理由] 名称変更

住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)

氏 名 インターナショナル・ビジネス・マシーンズ・コーポレーション